McDonald S, Wren C.

Consumer Search Ability, Price Dispersion and the Digital Divide.

Oxford Bulletin of Economics and Statistics (2016)

DOI: http://dx.doi.org/10.1111/obes.12151

CONSUMER SEARCH ABILITY, PRICE DISPERSION AND THE 'DIGITAL DIVIDE'

STEPHEN McDONALD* AND COLIN WREN

ECONOMICS, BUSINESS SCHOOL,
5 BARRACK ROAD, NEWCASTLE UNIVERSITY,
NEWCASTLE UPON TYNE, ENGLAND, UK, NE1 4SE.

ABSTRACT

The 'digital divide' in online activities is believed to arise from differences in Internet access, but this paper advances an alternative explanation that is related to consumer search ability. It argues that this leads to greater price dispersion, causing some consumers to be discriminated against. It analyses price data for the UK Internet motor insurance market, collecting data on 32,255 prices for 110 sub-markets, where differences in price dispersion across these by age, occupation and sex of the driver are argued to reflect differences in search ability. Allowing for price dispersion to also depend on the insurance risk, it finds greater price dispersion for consumers with weaker search abilities, i.e. older, unemployed, retired or female consumers. As this is not explained by alternative hypotheses, the paper concludes that improved Internet access alone will not close the 'digital divide'. The implication is that policymakers should address the online search abilities of individuals as well as Internet access.

KEYWORDS: Search ability, 'digital divide', price dispersion, Internet, insurance markets.

JEL CLASSIFICATION: L11, D83, D40.

*CORRESPONDING AUTHOR: stephen.mcdonald@ncl.ac.uk. Tel. +44 (0) 191 2081673.

ACKNOWLEDGEMENTS

Earlier versions of this paper were presented at the *Royal Economic Society* conference and to Economics staff at Newcastle University. The authors thank participants at these events, as well as Robert Hudson, Ginger Jin, Francis Kiraly, Claudio Piga, Steve Thompson and Mike Waterson for comments. The authors are grateful to referees of this journal, including the editor, Dr Debopam Bhattacharya, but remain responsible for the paper's content.

I. INTRODUCTION

The existence of a ‘digital divide’ is well documented, with older and poorer households less able to take economic advantage of the Internet due to their lower levels of online access (Hoffman and Novak, 2000; Chinn and Fairlie, 2006). The ‘digital divide’ will lead to price dispersion as the benefits of lower online search costs will not be distributed evenly. The evidence is that the uptake of the Internet reduced price dispersion (Brown and Goolsbee, 2002), and indeed Goldfarb and Prince (2008) find that the ‘digital divide’ no longer exists once Internet access is controlled for. This has meant that UK policy interest has centred on improving online access (ONS, 2012; OFT, 2007). Despite this, price dispersion remains pervasive in Internet markets for homogeneous goods and services (Haynes and Thompson, 2008; Baye *et al*, 2004), and this suggests that some consumers still do not take advantage of the lower online search costs and are being discriminated against. This paper explores an alternative explanation for price dispersion in online markets due to consumer search ability. It implies that improving Internet access alone is insufficient to close the ‘digital divide’.

To directly relate price dispersion to individual characteristics the paper takes a novel approach as it gathers price data for an online market by adopting assumed personal identities for different individual types. This is for the UK online motor insurance market in which the quotes are related to personal and car characteristics. As these individual types vary in their frequency of Internet use it is argued that they vary in their search ability. Further, through their experience of making quotes it is argued that the motor insurance firms learn this and that greater price dispersion will be observed for the consumers with weaker search abilities. The price data are collected from the leading UK motor insurance comparison website and from the leading sellers that choose not to list on these online aggregators. The quotes are potentially contractible and they are collected for 22 individual types by age, occupation and sex. This is for each of 5 car types, giving 110 sub-markets. Further, the data are collected at 12 monthly intervals, giving 1,320 price sets and a total of 32,255 price observations.

The use of insurance price data means that the insurance risk affects price dispersion, as well as the effectiveness of consumer search, although this is rarely considered in studies using insurance data (e.g. Dahlby and West, 1986). This is because in a sub-market, in which the individual characteristics are held constant, each insurance contract carries the same risk, so that in equilibrium a firm must sell more lower-priced contracts to get the same expected return as from selling higher-price contracts. This affects price dispersion and as higher risks mean that lower-priced contracts are less likely to be offered it also affects consumer search. In this paper, we regress the price data across sub-markets to examine the variation in price dispersion, controlling for this risk. Overall, we find that online price dispersion is greater for consumers that are older, unemployed, retired or female, indicating the existence of a 'digital divide'. These results are consistent with differences in consumer search ability, as measured by the frequency of Internet use, and they are not otherwise explained such as by unobserved heterogeneity. The implication is that improved Internet access alone will not be sufficient to close the 'digital divide', so that policy should also focus on online search ability.

In the next section, the basis for price dispersion at the sub-market level due to search behaviour is considered and the hypothesis set out. Section 3 outlines the data and Section 4 presents the results. These are discussed in Section 5, after which conclusions are drawn.

2. CONSUMER SEARCH AND PRICE DISPERSION

The relationship between consumer search behaviour and price dispersion is much explored in the literature and overall it finds that greater search will lead to less price dispersion (e.g.

Janssen and Moraga-Gonzàlez, 2004; Stahl, 1989).[1] This literature is characterised by many models of search behaviour, although not all of these are relevant to our particular setting. Thus, the main paper that relates to search in the motor insurance market by Dahlby and West (1986) assumes sequential consumer search behaviour, but it is inappropriate for an Internet market in which price comparison sites give multiple prices and consumer search costs are comparatively small. Further, the clearinghouse search model that is popular in the context of online markets (e.g. Morgan *et al*, 2006) is also inappropriate to our setting, as it supposes that the searching consumers discover all of the prices simultaneously, whereas in the UK motor insurance market the leading brands do not list on comparison websites.[2]

Sequential search is often viewed as the optimal search strategy, but non-sequential search, whereby consumers sample a fixed number of firms, may be optimal if there are fixed search costs or the consumers have knowledge of the market (De los Santos, 2008). Both of these features are apparent for the online motor insurance in which the cost of Internet access is fixed and insurance policies are renewed annually. In a market with non-sequential search price dispersion can arise either because individuals search at different fixed sample sizes (Burdett and Judd, 1983) or because they search randomly at the same fixed sample size, so that different prices are discovered by different individuals (Stigler, 1961).

The hypothesis that is explored in this paper is that an individual with a greater search ability faces lower price dispersion, as he or she is able to search more prices. It is explored at the sub-market level for the UK motor insurance industry. Consumer search ability is not measured directly, but rather price data are collected on motor insurance quotes for different individual types, where it is expected that consumers who are older, unemployed, retired or female have weaker search abilities on average. If this is the case, it suggests that a ‘digital

[1] Price dispersion may also arise from other factors, such as differences in marginal costs (Carlson and McAfee, 1983; Reinganum, 1979), but in insurance markets a substantial part of these costs will be the expected claim on a contract, which will again be constant at a sub-market level as the individuals carry identical risks.

[2] The clearinghouse model gives rise to the ‘gap’ measure of price dispersion, involving the difference between the lowest two prices (Baye *et al*, 2004), but this performs poorly when examined with our data below.

divide' exists related to consumer search ability. As a justification for this, Table 1 shows the frequency of Internet use for different types, where it is expected that a greater frequency of Internet use improves search ability by a 'learning-by-doing' argument (Arrow, 1962). Table 1 includes the time spent searching for prices or other online activities, and it holds Internet access constant by referring to users only. It shows that Internet usage declines with age and it is greater for males than females. Data on Internet usage are not available by occupation, but it is likely that higher-skilled jobs are associated with greater search abilities, since like Savage and Waldman (2009) they require a higher level of educational attainment.

The motor insurance market offers an ideal opportunity to examine the relationship between price dispersion and search ability as the quotes reflect the individual characteristics. As far as we are aware, only Brown and Goolsbee (2002) adopt a similar methodology, but they use life insurance data primarily to analyse the relationship between Internet access and the level of prices. Elsewhere, studies on online price dispersion examine differences across countries (Gatti and Kattuman, 2003), product characteristics (Chen and Scholten, 2003) or firm types (Pan *et al*, 2002), but they do not explicitly investigate search. Bayliss and Perloff (2002) attribute online price dispersion to search by ruling out non-search explanations. The search abilities are also not directly observed by firms, but as they offer insurance repeatedly to individual types that differ by age, occupation and sex, it is reasonable that they will learn the types that search with most intensity and it is these characteristics that we focus on. The pricing strategies are embedded in the firms' computer programs, but these are commercially confidential and would have to be obtained for all firms to determine price dispersion.

Finally, the price quote data are collected by four ages, 25, 40, 55 and 70 years, which represents a good age spread, and for each of these for the occupations of factory worker and computer consultant, as well as an unemployed and a retired person.[3] For realism, a 70-year

[3] The motor insurers that do not list on *confused.com* do not permit search at this level, so that 'manual' or

old is retired only, but 25- and 40-year olds are not retired, so there are eleven age-occupation combinations. The data are collected for both males and females, giving 22 consumer types. According to the above, older individuals and females are expected to have weaker search abilities, and a factory worker to be less skilled than a computer consultant. A retired or an unemployed person is likely to be less skilled than the employed occupations, although the retired are expected to be more skilled than the unemployed, although they are heterogeneous and include the so-called 'silver-surfers' as well as the inactive. The data were collected at twelve monthly intervals over 2006-07, which is the period referred to in Table 1. While the frequency of online use has increased since then, Table 1 shows that the relative usage pattern has remained much the same, while the major UK motor insurers continue not to be listed on comparison websites, so that consumer search ability remains an issue.

3. THE MOTOR INSURANCE MARKET AND DATA

The insurance premium quotes are collected from the main UK motor insurance comparison website, *confused.com*, and directly from the websites of the four leading motor insurers that choose not to list on comparison websites.[4] Consumers enter their details at *confused.com*, which then visits the websites of quoted insurers and reports offered prices, ranking them from lowest to highest, along with the policy details and brand information. It generates its revenue from sales generated from click-through or telesales made using a reference number. At the time of data collection *confused.com* claimed to cover 94% of UK motor insurance firms and it had a two-thirds share of motor insurance policies sold in the UK via an online aggregator (*Financial Times*, 2007). The leading brands that are not listed on comparison sites account for 30% of the market, and this is still the case for the heavily advertised brands

'manufacturing and engineering' and 'professional' or 'scientific and technical' are instead used for these.
[4] The four unlisted sellers are *Churchill*, *Direct Line*, *Esure* and *Tesco*.

(McDonald and Wren, 2012). It means that the consumer cannot discover all of the market prices by consulting a single comparison site. It is reasonable that the consumers wishing to sample more prices will go to the sites of at least one of these four leading brands.

The premium quotes, which are contractible, were collected on a consistent basis for a fully comprehensive motor insurance cover. This represents 90% of the UK market. Apart from the driver characteristics, which are used to proxy the consumer search ability, the other contract characteristics are specified to ensure a homogeneous good for each car type. These reflect average market conditions and include an annual mileage of 9,000 miles, a 5-year no-claims bonus, a no-claims bonus protection and a business, commuting and social use. Other policy details are held constant, including the marital status of the driver, type of licence, motoring convictions, car alarm and so on, so that apart from some minor policy details that cannot be controlled for, for all intents and purposes the contracts are homogeneous.[5]

The legal fees cover is set to zero as few firms include this as standard, but a courtesy car is requested. About 90% of the policy quotes include a 'compulsory excess', which is the deductible in the event of a claim. It is specified by the insurer, on which details are collected. The 'voluntary excess' is specified by the policyholder and so set to zero. Motor insurance varies by location and a single address is selected. This is a reasonably affluent suburb of Newcastle-upon-Tyne, which has a mean house price close to that of the national average.[6]

As regards the car type, nine market segments are identified by the *UK Society of Motor Manufacturers and Traders*, so that to be representative of this market, and to ensure the generality of the results, the car type is chosen as the leading model in five of the major market segments, based on UK sales (other car segments are either small markets or similar

[5] They include a married individual with no children, who has held a licence since age 18 years, a resident homeowner with no motoring convictions or penalties in last 5 years and who has no other cars or additional drivers. The car has no alarm, mobilizer or tracker and no modifications.
[6] This is a street in the Gosforth postcode district. Forty streets were randomly sampled, and a street chosen with a standard deviation of quoted prices across firms closest to the mean deviation. 25 firms offered prices for some or all of the 40 streets, of which 13 offered the same price at each location. Of the remainder, the greatest standard deviation across locations for any firm was 14.7% of the mean premium, which is small.

to the ‘upper-medium car segment’). These are the Ford Fiesta Encore, Ford Focus Zetec, Vauxhall Vectra CD 16V, BMW 525i and MG MGF. Details of these are given in Table 2, where at the time of data collection each car is the median car age (SMMT, 2006). Table 2 shows the value and sales of each car and it gives an index of risk. According to this index, the BMW is a higher-risk model and the Fiesta and Focus cars are low-risk models, although in practice the insurance risk will also depend on the driver characteristics.

The price data relate to 110 sub-markets (i.e. 22 individual and 5 car types), and these were collected over a one-year period at monthly intervals from February 2006 to January 2007, giving 1,320 price sets. To minimise the effect of updating, the insurance premium quotes were collected from the websites over a weekend for each month. A firm may offer more than one price in a sub-market, but the data records the details of a firm’s lowest quoted price only. Further, a seller could have multiple brands, but the firms conceal ownership (see McDonald and Wren, 2013), so that these are treated as different sellers. The main methods of selling motor insurance are represented in the data: an underwriter selling insurance direct to the consumer; through an exclusive relationship with a firm or firms; or through a broker that deals with a panel of underwriters. Overall, there are 32,255 observations on prices by 41 firms in the dataset. The mean number of prices in a set is equal to 24 (i.e. 32,255 / 1,320), but varying between 11 and 31 prices. Entry and exit mean that the average number of firms quoting in each month is between 31 and 36 firms, with 27 firms quoting in every month.

Descriptive statistics for the insurance premiums across the individual types are given in Table 3. This is for the 1,320 price sets, disaggregated by the five car types. The cars have different mean premiums that reflect the risks shown in Table 2. Table 3 reveals that there is consistent pattern across the individual types, with higher premiums tending to paid by those that have higher accident rates or more severe accidents, e.g. 25 year-old males. However, the evidence for the effect of search ability on price dispersion is weak. Measured by the coefficient of variation of the prices, Table 3 shows that this is reasonably invariant across the individual and car types, so that there is no clear pattern (although it is higher for some high-risk individual-car combinations, such as a 25-year old driving the MGF car). However, it could be that the individual types embody both search and risk characteristics.

4. THE MODEL AND RESULTS

The model is regressed across the 1,320 price sets, which comprises data on 110 sub-markets (j = 22 individual and k = 5 car types) for each of t = 12 months. It is specified as:

$$y_{jkt} = \alpha + \beta_1 s_j + \beta_2 r_{jkt} + \beta_3 r_{jkt}^2 + \beta_4 z_{jkt} + \delta_k + \delta_t + \varepsilon_{jkt}, \qquad (1)$$

where y_{jkt} is the dispersion of prices for a price set, s_j are the search intensity terms, r_{jkt} is the insurance risk of the price set, z_{jkt} are controls that capture other features of the UK online motor insurance market and δ_k and δ_t are fixed effects for the cars and months. As outlined above, the hypothesis that is examined is that greater online search ability leads to lower price dispersion, where this ability is captured by the search intensity s_j in equation (1), and proxied by dummies for the age, occupation and sex of the driver as in Table 3. These proxies could potentially capture other explanations for price dispersion, such as differences in the search propensity of individuals or some other form of heterogeneity. It is not possible to control for all these, some of which are unobservable, so that alternative explanations for the regression results are considered in Section 5. We now discuss the specification in equation (1) and the measurement of these other terms, before setting out the estimation results. Further details of the measurement of these other variables are given in Table 4.

4.1. THE VARIABLES

The prices are the gross insurance premiums, but as these can be relatively high in a price set then these prices are logged. The price dispersion y_{jkt} in (1) is measured using the coefficient of variation (*CV*) of the prices for each price set, which has the advantage of weighting the standard deviation of prices by the mean price of the price set. However, as we use the mean premium to capture the insurance risk (see below), then alternative measures of dispersion are also considered, as set out in Table 4. These are the standard deviation, range and a measure that allows for the compulsory excess amount (*CVADJ*), as in the table notes.

As mentioned in Section 2, price dispersion may be affected by the insurance risk of a price set, reflecting both the driver and car characteristics. This is because higher risk makes lower prices relatively less attractive, as the firm must sell more lower-priced contracts to get the same expected return as from higher-price contracts, suggesting that within a price set the prices are less dispersed.[7] The risk also affects the effectiveness of consumer search since for any given level of search a higher risk means that the lower prices are less likely to be posted. It is not possible to interact the risk and search terms as it makes the search terms difficult to interpret, but as the risk has opposing effects on price dispersion and search r_{jkt} is included in a quadratic form in equation (1), where $\beta_2 > 0$ and $\beta_3 < 0$. It means higher risk reduces the price dispersion at a decreasing rate. The quadratic form captures the effect of risk on search intensity and the sensitivity of the results to this is explored below.

To measure the sub-market risk r_{jkt} a risk index is constructed that is based on the UK motor insurance claims data (see Table 4). However, as this index is not available for each of age, occupation, sex and car type, we also use the mean premium of the price set to measure

[7] As a simple illustration, suppose insurance contracts are on offer at two prices: £10 and £15. If the expected claim on a contract is £5 the ratio of expected returns is (£10 - £5) / (£15 - £5) = 0.5, but if the risk increases so that the expected claim is £7.50 then the ratio of expected returns falls to a third, so that more insurance policies must be sold at £10 to get the same expected return from selling policies at £15.

the risk on the basis that the more riskier sub-markets will face higher premiums on average. Table 3 shows that higher premiums are on average paid by the individuals that have higher accident rates and more severe accidents (e.g. the young and males) and by car type the mean premium is correlated with the risk ratings in Table 2. The advantage of the mean premium is that it is available for each individual and car, but a difficulty is that it is contemporaneous with the sub-market price dispersion. It is addressed by considering the initial mean premium (Table 4), which measures the mean premium for each sub-market at February 2006 and uses this in a regression of (1) for the other eleven months from March 2006 to January 2007.

The z_{jkt} in equation (1) are controls for other effects that determine price dispersion at a price set level. First, a term is included for the mean compulsory excess. As this lowers the risk a positive sign is expected. Second, the number of sub-market drivers is included. As a larger number of drivers makes lower prices more feasible a positive sign is expected. Since these data are not directly available, it is measured by the number of UK drivers by age or sex, and multiplied by the relative size of the five car markets. Third, the strategic behaviour of firms is allowed for by including terms for the number of firms that declare themselves either as quoting only for given individual or car types.[8] The signs on these depend on whether the individual or car types are towards the middle or extreme of the price distribution. Finally, the z_{jkt} include dummies for each firm that take a value of unity if the firm quotes in the price set. These dummies may capture differences in costs, in consumer preferences for quality if some firms have a reputation for easier claims settlement, or in capacity constraints that may restrict the issuance of motor insurance policies (see Gron, 1994a).

4.2. ESTIMATION RESULTS

[8] Table 4 shows that at least two firms never quote for some individual types.

The results from estimating equation (1) are presented in Table 5. The first three columns of this table examine the sensitivity of the results to the different risk measures. Column I uses the index constructed from the claims data, column II uses the mean premium across all price sets and column III uses the initial mean premium at February 2006 regressing the equation across 1,210 price sets. In terms of the model fit and sign and significance of the estimates, Column III gives the best results, so that this is discussed below. Column II produces similar results to column IIII, but column I gives the weakest results with a significant negative sign on the number of drivers that could be related to the nature of the claims data. Columns IV and V use alternative price dispersion measures, but give virtually identical results to column III. Columns VI and VII show that female term is insignificant or wrongly signed when the risk term is linear or omitted altogether, which gives support for the model specification.

Focusing on column III in Table 5, the risk terms have the expected signs, and when they are evaluated at the mean values in Table 4 risk has a negative effect on price dispersion. Column III shows that a greater excess leads to lower price dispersion, signalling lower risk, while greater sub-market size (more drivers) leads to greater dispersion. The opt-outs show that firms choose not to quote in sub-markets for the individuals at the extremes of the price distribution, suggesting that they avoid risky individual types. By car type they do not quote for sub-markets towards the middle of the distribution, suggesting more-specialised markets. As regards the fixed effects, only four of these firm dummies are significant that may reflect the competitive nature of the UK online motor insurance market (Competition Commission, 2013). The car dummies are each significant, while as a group the monthly time dummies are insignificant, which is possibly because the number of firms that quote in the motor insurance market as a whole varies little over the period of a single year.[9]

[9] There is less variation in price dispersion over time, which is consistent with each firm offering a similar price in each sub-market over time. It may mean that the risk and error terms are correlated from unmeasured effects. However, as the prices are logged then omitted terms are likely to have a multiplicative effect on the variance of the prices, which is controlled for by measuring price dispersion y_{jkt} using the coefficient of variation since CV

As regards the search terms, the base case for the individual terms in Table 5 is a 25-year old, unemployed male. The estimates show that price dispersion is greater for the old compared to the young, greater for the unemployed and retired compared to the employed, and greater for females relative to males. There is no significant difference between the 25- and 40-year olds, which is consistent with the frequency of Internet use in Table 1, although price dispersion is greater for the 55-year olds and greater again for the 70-year olds, each at the 1% significance level. There is also no significant difference between the occupations of factory worker and computer consultant and this is discussed below. The retired face less price dispersion than the unemployed, but more dispersion than the employed, again each at the 1% significance level. The estimates on the search terms are interpreted below.

4.3. FURTHER EXPLORATIONS

Before interpreting the results, further explorations were made to examine the sensitivity of the estimates. First, the price quotes consist of both a premium and a compulsory excess, but as the excess has a negative effect on the premium then it may affect price dispersion.[10] To examine this, the excess-adjusted measure of price dispersion (*CVADJ*) is instead used (Table 4). This is the certainty equivalent of each premium-excess price quote based on a constant relative risk aversion utility function (McDonald, 2010). This was constructed under a range of assumptions about the consumer income and risk aversion, but the estimates were robust to this, possibly because the compulsory excess is already included in the estimations in Table 5. Column VIII of Table 6 gives a representative regression, and it shows that price dispersion now increases strictly with age, but overall the results are like those column III of Table 5.

Second, the insurance premiums vary by location and so the residential address of the

is homogeneous of degree zero, so that CV(ax) = CV(x), where a is an unmeasured effect.

[10] Regressing the premium on the excess and dummies for the car type gives a significant negative coefficient of 0.06, so that on average a £1 reduction in the excess amount increases the premium by 6p.

driver was examined. This involved collecting data for four other locations, but for the Focus car only. Two locations were selected at similar suburban locations elsewhere in England (i.e. Bristol and Reading) and two in the same geographical area (i.e. Byker and Hexham), where Byker is a deprived urban area of Newcastle and Hexham a prosperous rural market town. These give a further 244 price sets, each giving an observation on price dispersion in addition to the observations on Gosforth used above. The results from estimating equation (1) for the Focus car only with the search terms s_j in spline form for these locations are given in column IX. This regression performs less well and the estimate on female is negative and significant. However, like column VII of Table 5, the risk estimates are insignificant, suggesting that the individual terms may now capture this effect. Overall, it shows the importance of controlling for risk, for which a regression across car types seems important. Nevertheless, it does show a broadly consistent pattern of results across the five geographical locations.

Finally, instead of regressing (1) across price sets, the data were regressed across the $i = 1, 2, 3, \ldots, 32{,}255$ prices, and the hypothesis examined by observing whether the errors are less dispersed for individuals that are believed to have greater search ability. In measuring risk it is not possible to include the mean premium of a price set, since the regression returns an estimate of unity for this, so that the minimum premium of the relevant price set p_0 is used, which is measured for each individual, car and month. This is exogenous and it is also likely to be correlated with risk. The minimum premium $\underline{p_{jkt}}$ is deducted directly from the price, so that the estimating equation is as follows, where the error variance depends on search:

$$\ln (p_i / \underline{p_{jkt}}) = a + b_1 s_j + b_2 z_{jkt} + \delta_k + \delta_t + \varepsilon_i, \qquad (2)$$

$$\text{where } \varepsilon_i \sim \mathrm{N}(0, \sigma_i^2) \quad \text{and} \quad \sigma_i^2 = a' + b' s_j.$$

Column X of Table 6 presents the result from regressing (2) using maximum likelihood. The estimates on the error variance are given in the second column and they are consistent with column III of Table 5, i.e. greater price dispersion for the old, unemployed and females, but the retired no longer show significantly less price dispersion than the employed.

5. DISCUSSION OF THE RESULTS

Overall, the empirical results show that price dispersion is greater for those individuals who are in general older, unemployed, retired or female. As these types have a lower frequency of Internet use then it is consistent with the hypothesis that weaker consumer search ability leads to greater price dispersion. However, as the findings may potentially be consistent with other explanations then before concluding it is necessary to consider these.

First, the results suggest that the insurance risk affects price dispersion, and as noted above it may also affect the relationship between consumer search and price dispersion. This is because higher risk means that lower prices are less likely to be posted, making search less effective. While the risk terms are included in the empirical model in a quadratic form and different risk measures are considered, we are unable to interact the risk and search terms, but with the implication that the individual terms may just capture the insurance risk. However, the results for the individual terms are not consistent with a risk explanation. This is because price dispersion is greater for older and female drivers, who are known to be less risky. The estimates on the individual terms can also differ significantly between the occupations where there is likely to be little or no variation in risk, so that we do not believe that this provides an explanation for the pattern of estimates obtained on the individual terms.

A second possible explanation is that the results reflect capacity constraints in the UK motor insurance industry, perhaps due to insurers' restricted access to finance, leading them to offer higher prices in some sub-markets. Of course, this argument may primarily relate to the level of prices rather than price dispersion and it is controlled for in the empirical analysis by using the coefficient of variation. Further, to the extent that capacity constraints matter it is expected that they will be most apparent in the riskier sub-markets, but it was noted above that the pattern of estimates is not consistent with a risk explanation. In the case of the UK motor insurance industry it was making underwriting losses at the time of data collection, but it was able to more than recoup these losses through its investments (Meeson, 2009).

As a further possible explanation for the results it could be that the individuals differ in their unobserved characteristics and this affects their online search behaviour. Thus, some individual types may engage in lower search levels as they have a higher proportion of loyal consumers, e.g. older consumers have a more-established relationship with an insurer and are more reluctant to switch insurers. However, the prices posted online reflect the individuals who actually engage in search, i.e. non-loyal consumers, while McDonald and Wren (2012) find that consumer loyalty does not increase with age, so that the results do not support this. The individuals may also differ in their online access or ability to search using other channels, such as through a broker or by telephone. However, if search by these alternative channels matters then online prices will depend on offline behaviour, but it is well known that firms discriminate between Internet and non-Internet users (Brynjolfsson and Smith, 2000).[11] As the product characteristics are held constant, while firm dummies are included, the results are not believed to reflect unobserved heterogeneity in the individual characteristics.

Finally, it is possible that the results just reflect differences in the propensity to search of individuals. Indeed, the results for younger drivers are consistent with this as these types have lower incomes and have a greater motivation to search, but face less price dispersion. However, while search propensity could be present, and may explain why price dispersion is

[11] Brown and Goolsbee (2002) find that the level of Internet access affects offline prices, although this is for early Internet uptake in which consumers search online but are required to purchase offline.

more pronounced for age compared to occupation, it does not fully account for the results.[12] This is because price dispersion is greater for the unemployed and retired, while relative to the employed these individual types each have lower incomes and an opportunity cost of time. Likewise, there are significant differences in price dispersion in the retired between those aged 55 and 70-years, even though these are likely to have similar opportunity costs of time. Price dispersion is significantly lower for 55-year old retirees, when the opposite is expected if these individuals are able to afford to retire early and search propensity is important.

6. CONCLUSIONS

Price dispersion remains pervasive in Internet markets and this paper has sought to explain this in terms of differences in consumer online search ability. Search ability is proxied by the individual characteristics associated with the frequency of Internet use, and the paper collects price data from the UK online motor insurance market that are directly related to these types. The empirical results show that price dispersion is greater for the individual types that have a lower frequency of Internet use, including older, unemployed, retired and female consumers. These results are consistent with a lower level of search ability, while they cannot otherwise be explained, suggesting that ability is important in explaining online price dispersion. Given that this discrimination is apparent for older, unemployed, retired and female consumers then it points to the existence of a 'digital divide' related to the search skills of individuals. This not only explains why a 'digital divide' persists, despite the efforts to improve Internet access, but it suggests that the online search skills of individuals are important and that policymakers should address these Internet search skills in order to close the 'digital divide'.

[12] It could also account for some other features of the results. For example, we find that there is no significant difference between the factory worker and computer consultant. This could be because the consultants have greater search ability, but higher incomes, and so a lower search propensity. In general, we find no difference between the 25- and 40-year olds, although this could be that these ages have similar search abilities (Table 1).

TABLE 1: FREQUENCY OF INTERNET USE

Computer use in previous 3 months (%)		Daily	Weekly	Monthly
Age:	16-24 years	63 (82)	15 (12)	7 (2)
	25-44	62 (83)	16 (10)	4 (2)
	45-54	56 (71)	13 (14)	5 (2)
	55-64	36 (63)	17 (13)	5 (2)
	65 plus	9 (29)	8 (11)	3 (4)
Sex:	Male	73 (81)	20 (19)	7 (4)
	Female	65 (76)	24 (22)	11 (5)

Notes: Figures for 2006; they do not sum to 100. Sex on a different basis to age (i.e. 'every day or almost every day'). Figures in parentheses are 2012 for age and 2010 for sex.
Source: 'Internet Access: Households and Individuals', *Statistical Bulletin*, Office for National Statistics, London, various years.

TABLE 2: THE CAR TYPES

Make and model	Market segment	Market share	Engine size	Car value	Total sales	Risk rating
Ford Fiesta Encore	Super-mini	13.3%	1299cc	£1,595	91,783	4
Ford Focus Zetec	Lower-medium	17.3%	1596cc	£3,350	114,512	5
Vauxhall Vectra CD 16V	Upper-medium	14.3%	1998cc	£3,120	70,704	12
BMW 525i	Executive	12.9%	2494cc	£8,065	13,443	16
MG MGF	Sports	8.6%	1796cc	£4,320	5,766	12

Notes. Prices at 2006 and sales at 2000. Advisory risk rating of the Association of British Insurers based on damage and parts costs, repair times, new car values, body shells, performance and security. Risk is in range 1 and 20, where 1 is lowest.
Sources: *www.parkers.co.uk* (February 2006) and SMMT (2006).

TABLE 3: PREMIUM BY INDIVIDUAL AND CAR TYPE

Individual type		Fiesta	Focus	Vectra	BMW	MGF
Age:	25 years	373	389	570	905	692
		(0.12)	(0.12)	(0.14)	(0.22)	(0.25)
	40 years	269	281	403	526	465
		(0.13)	(0.13)	(0.12)	(0.11)	(0.22)
	55 years	212	223	306	400	371
		(0.18)	(0.16)	(0.16)	(0.13)	(0.28)
	70 years	223	232	325	414	384
		(0.21)	(0.18)	(0.18)	(0.16)	(0.26)
Occupation:	Unemployed	299	316	455	653	577
		(0.14)	(0.13)	(0.14)	(0.16)	(0.42)
	Factory	281	293	417	600	486
		(0.14)	(0.14)	(0.15)	(0.16)	(0.17)
	Consultant	276	286	410	581	473
		(0.14)	(0.13)	(0.14)	(0.15)	(0.17)
	Retired	215	225	312	402	371
		(0.21)	(0.18)	(0.18)	(0.15)	(0.24)
Sex:	Male	278	290	412	584	495
		(0.14)	(0.14)	(0.14)	(0.15)	(0.25)
	Female	267	280	400	562	448
		(0.17)	(0.14)	(0.15)	(0.16)	(0.25)
All individuals		273	285	406	573	486
		(0.15)	(0.14)	(0.15)	(0.15)	(0.25)

Notes: Mean premiums rounded to nearest £ and coefficient of variation (CV) in parentheses. Each is calculated as the mean across the relevant price sets. Compulsory excess not included.

TABLE 4: DESCRIPTIVE STATISTICS

Variable	Mean	Standard deviation	Minimum	Maximum
Dependent variable (y_{jkt}):				
Coefficient of variation (*CV*)	0.027	0.008	0.015	0.059
Standard deviation (*SD*)	0.157	0.049	0.085	0.350
Range (*RANGE*)	0.701	0.264	0.300	1.775
Excess-adjusted CV (*CVADJ*)	0.024	0.008	0.013	0.056
Risk-adjusted price (ln (p_i / $\underline{p}_{jkt}$))	0.320	0.201	0	2.236
Risk terms (r_{jkt}):				
Claims index	0.754	0.507	0.170	2.360
Claims index^2	0.826	0.114	0.029	5.570
Mean premium	-0.054	0.271	-0.469	0.616
Mean premium^2	0.076	0.067	0.48 x 10^{-6}	0.380
Initial mean premium	-0.032	0.270	-0.396	0.612
Initial mean premium^2	0.074	0.071	0.48 x 10^{-6}	0.375
Other terms (z_{jkt}):				
Compulsory excess	4.994	0.170	4.476	5.399
Number of drivers	10.527	1.210	7.968	11.981
Individual opt-out	5.557	1.842	2	9
Car opt-out	1.959	3.241	0	9

Notes: Statistics calculated across 1,313 price sets (seven price sets dropped due to errors in data collection), except that ln (p_i / $\underline{p}_{jkt}$) is across 32,255 prices. Natural logs of premium and excess (£'s). Range is difference between maximum and minimum price of a price set, divided by mean premium, with natural logs taken. Excess-adjusted CV is certainty equivalent premium, assuming income of £20,000 and relative risk aversion index of 1.5. Claims index is actual claims by age and sex, weighted by car risk rating relative to 20, shown in Table 2. Initial mean premium measured at February 2006, but March 2006 for missing cases. Initial mean premium and mean premium are centred about car type (i.e. mean log price for a price set minus the mean log price for the car type). Number of drivers is at 2006 and expressed in logs.
Source: Authors' own data, except claims (Association of British Insurers) and driver numbers (DVLA, UK).

TABLE 5: REGRESSION RESULTS

Column:	I	II	III	IV	V	VI	VII
Dependent variable (y_{jkt}):	*CV*	*CV*	*CV*	*SD*	*RANGE*	*CV*	*CV*
Constant	200.47***	-110.33	-575.40***	-329.74***	-216.19***	-122.61***	-216.19***
Individual terms (s_j):							
40 years	15.50***	9.96***	-2.25	-1.17	-0.82	13.46***	15.35***
55 years	21.63***	29.38***	28.69***	16.57***	9.57***	30.96***	17.56***
70 years	17.20***	38.09***	62.29***	36.08***	21.93***	41.48***	2.30
Factory worker	-12.63***	-10.82***	-8.80***	-5.61***	-3.13***	-14.53***	-11.40***
Computer consultant	-13.51***	-11.08***	-8.31***	-5.33***	-2.90***	-14.40***	-12.20***
Retired	-11.17***	-8.29***	-5.08***	-3.48***	-2.09***	-12.31***	-10.03***
Female	-0.65	0.61	6.01***	3.44***	2.51***	1.31	-4.26***
Risk terms (r_{jkt}):							
Claims index	20.30***	-	-	-	-	-	-
Claims index^2	-4.19***	-	-	-	-	-	-
Mean premium	-	38.96***	-	-	-	-	-
Mean premium^2	-	-28.84***	-	-	-	-	-
Initial mean premium	-	-	61.77***	37.59***	22.81***	36.59***	-
Initial mean premium^2	-	-	-51.72***	-28.84***	-15.54***	-	-
Other terms (z_{jkt}):							
Compulsory excess	-7.33***	-5.27***	-7.54***	-4.73***	-2.00***	-4.65***	-5.84***
Number of drivers	-12.62***	14.05**	55.42***	31.94***	20.16***	15.18**	-25.42***
Individual opt-out	-2.18***	-2.65***	-2.14***	-1.30***	-0.74***	-3.41***	-1.81***
Car opt-out	6.44***	4.52***	3.40***	2.37***	0.13	4.23***	7.61***
F	84.8***	88.3***	88.2***	76.8***	60.8***	87.5***	89.2***
R^2	0.69	0.71	0.73	0.73	0.66	0.70	0.68
No. of observations	1,313[a]	1,313[a]	1,210	1,210	1,210	1,210	1,210

Notes: OLS estimation of equation (1) with clustered standard errors for sub-market correlation over time. Car, month and firm fixed effects included throughout, but not shown. *CV* estimates multiplied by 1,000, *SD* by 100 and *RANGE* by 10. Variables described in Table 4. [a] Seven price sets dropped due to errors in data collection. *** Significant at the 1%, ** = 5% and * = 10% level.

TABLE 6: FURTHER EXPLORATIONS

Column:	VIII	IX					X	
Dependent variable:	*CVADJ*	*CV*					ln (p_i / p_{jkt})	
Constant	200.47***	605.95					40.41	30.31***
Individual terms (s_j):		Gosforth	Byker	Hexham	Bristol	Reading	ln (p_i / p_{jkt})	σ_i^2
40 years	12.92***	19.16***	1.57	1.15***	21.70***	18.66***	1.57	1.15***
55 years	39.14***	17.17***	8.34***	10.32***	11.25**	12.02**	8.34***	10.32***
70 years	47.02***	-9.77	7.62***	7.95***	-17.22**	-17.30**	7.62***	7.95***
Factory worker	-9.09***	-8.59***	-5.09***	-6.84***	-2.67	-5.63**	-5.09***	-6.84***
Computer consultant	-8.64***	-9.13***	-4.86***	-7.52***	-5.36*	-5.87**	-4.86***	-7.52***
Retired	-5.34***	-7.68***	-5.57***	-6.90***	-3.77	-5.70**	-5.57***	-6.90***
Female	3.52**	-7.82***	-0.61	0.78***	-10.76***	-8.69***	-0.61	0.78***
Risk terms (r_{jkt}):								
Initial mean premium	71.32***	6.42					-	-
Initial mean premium^2	-53.19***	-1.26					-	-
Other terms (z_{jkt}):								
Compulsory excess	-7.47***	-6.71**					-0.01***	-
Number of drivers	16.97**	-41.83***					-1.40	-
Individual opt-out	-2.05***	-1.85***					-0.96***	-
Car opt-out	3.20***	-					8.64***	-
χ^2 [F]	[66.9***]	[80.4***]					2,808***	
LogL [R^2]	[0.70]	[0.63]					11,314.9	
No. of observations	1,210	483[a]					32,255	

Notes: Column VIII is equation (1) using *CVADJ* as dependent variable, column IX estimates equation (1) for Focus car with additional data for location in spline form, and column X regresses equation (2). Estimates multiplied by 1,000 in columns VII and IX and by 100 in column X. Variables described in Table 4. [a] Three price sets dropped due to errors in collection. *** significant at the 1% ** = 5% and * = 10% level.

REFERENCES

Arrow, K. J. (1962), 'The Economic Implications of Learning By Doing', *Review of Economic Studies*, 29, 155-73.

Baye, M., Morgan, J. and Scholten, P. (2004), 'Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site', *Journal of Industrial Economics*, 52, 463-96.

Baylis, K. and Perloff, J. (2002), 'Price Dispersion on the Internet: Good Firms and Bad Firms', *Review of Industrial Organization*, 21, 305-24.

Brown, J. and Goolsbee, A. (2002), 'Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry', *Journal of Political Economy*, 110, 481-507.

Brynjolfsson, E. and Smith, M. (2000), 'Frictionless Commerce? A Comparison of Internet and Conventional Retailers', *Management Science,* 46, 563-85

Burdett, K. and Judd, K. (1983), Equilibrium Price Dispersion, *Econometrica*, 51, 955-69.

Carlson, J. and McAfee, R. P. (1983), 'Discrete Equilibrium Price Dispersion', *Journal of Political Economy*, 91, 480-93

Chen, J. and Scholten, P. (2003) 'Price Dispersion, Product Characteristics, and Firm Behaviors: Stylized Facts from Shopper.com', *Advances in Applied Microeconomics*, 143-64.

Chinn, M. and Fairlie, R. (2006), 'The Determinants of the Global Divide: A Cross-Country Analysis of Computer and Internet Penetration', *Oxford Economic Papers,* 59, 16-44.

Competition Commission (2013), *Private Motor Insurance Market Investigation: Provisional Findings Report*, Competition Commission, London, December.

Dahlby, B. and West, D. S. (1986), 'Price Dispersion in an Automobile Insurance Market', *Journal of Political Economy,* 94, 418-38.

Data Monitor (2008), *UK Private Motor Insurance 2007,* London, 7th February 2008.

De los Santos, B. (2008), 'Consumer Search on the Internet', *Working Paper 08-15*, Kelley

School of Business, Indiana University.

Financial Times (2007), 'Admiral in Talks over Confused.com Sale', *Financial Times*, London, 25th May 2007.

Gatti, J. R. J. and Kattuman, P. (2003), 'Online Price Dispersion Within and Between Seven European Countries', *Advances in Applied Microeconomics*, 107-41.

Goldfarb, A. and Prince, J. (2008), 'Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide', *Information Economics and Policy,* 20, 2-15.

Gron, A. (1994a), 'Evidence of Capacity Constraints in Insurance Markets', *The Journal of Law and Economics,* October, 349-77.

Gron, A. (1994b), 'Capacity Constraints and Cycles in Property-Casualty Insurance Markets', *RAND Journal of Economics,* 25.1, 110-27.

Haynes, M. and Thompson, S. (2008), 'Price, Price Dispersion and Number of Sellers at a Low Entry Cost Shopbot', *International Journal of Industrial Organisation*, 26, 459-72.

Hoffman, D. L. and Novak, T. P. (2000), 'The Growing Digital Divide: Implications for an Open Research Agenda', in E. Brynjolfsson and B. Kahin (eds), *Understanding the Digital Economy: Data, Tools Research,* MIT Press, Cambridge, MA, pp. 245-60.

Janssen, M. and Moraga-Gonzàlez, J. L. (2004), 'Strategic Pricing, Consumer Search and the Number of Firms', *Review of Economic Studies,* 71, 1089-118.

McDonald, S. (2010), 'Consumer Search Costs, Pricing Strategies and Price Dispersion: The Internet Motor Insurance Market', *PhD Thesis*, Newcastle University, UK.

McDonald, S. and Wren, C. (2012), 'Informative Brand Advertising and Pricing Strategies in Internet Markets with Heterogeneous Consumer Search', *International Journal of the Economics of Business*, 19, 55-69.

McDonald, S. and Wren, C. (2013), 'Multiple Price Posting and Consumer Search Obfuscation: Evidence from an Online Market', *International Industrial Organization* *conference*, Boston, Mass, May.

McKenna, C. J. (1986), *The Economics of Uncertainty*, Wheatsheaf Books, Brighton.

Meeson, L. (2009), 'Motor Insurance Industry Set to Make £1bn Underwriting Loss, Claims Deloitte', *Insurance Age*, December.

Morgan, J., Orzen, H. and Sefton, M. (2006), 'An Experimental Study of Price Dispersion', *Games and Economic Behaviour*, 54, 134-58.

Morgan, P. and Manning, R. (1985), 'Optimal Search', *Econometrica*, 53, 923-44.

OFT (2007), *Internet Shopping, Economic Literature Review*, Report for the Office of Fair Trading, by Europe Economics, Annex F, Office of Fair Trading, London.

ONS (2012), Internet Access: Households and Individual, 2012', *Statistical Bulletin*, Office for National Statistics, London, 24th August.

Pan, X., Shankar, V. and Ratchford, B. T. (2002), 'Price Competition Between Pure Play Versus Bricks-And-Clicks E-Tailers: Analytical Model and Empirical Analysis', *The Economics of the Internet and E-Commerce*, 29-61.

Reinganum, J. (1979), 'A Simple Model of Equilibrium Price Dispersion', *Journal of Political Economy*, 87, 851-8.

Savage, S. and Waldman, D. (2009), 'Ability, Location and Household Demand for Internet Bandwidth', *International Journal of Industrial Organization*, 27, 166-74.

SMMT (2006), *Motor Industry Facts - 2006*, Society of Motor Manufacturers and Traders, London.

Stahl, D. O. (1989), 'Oligopolistic Pricing with Sequential Consumer Search', *American Economic Review*, 79, 700-12.

Stigler, G. J. (1961), 'The Economics of Information', *Journal of Political Economy*, 69, 213-25.